Search for Boosted Top Quarks at CDF

Outline

1. Introduction and Motivation
2. Data Selection & Jet Calibration
3. Boosted Top Signals
4. Results
5. Conclusions

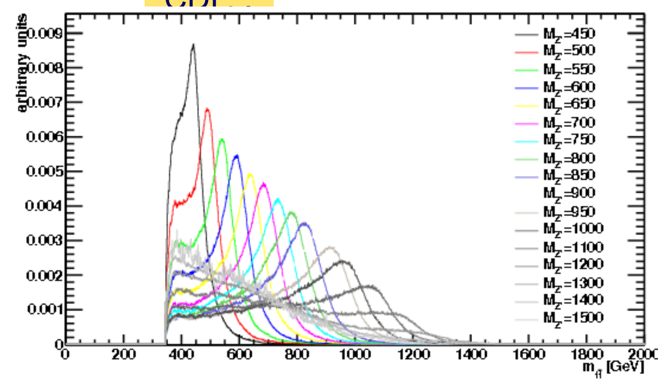
Representing the CDF Collaboration

Key Players:
Raz Alon, Gilad Perez & Ehud Duchovni
Weizmann Institute of Science
&
Pekka K. Sinervo, FRSC
University of Toronto

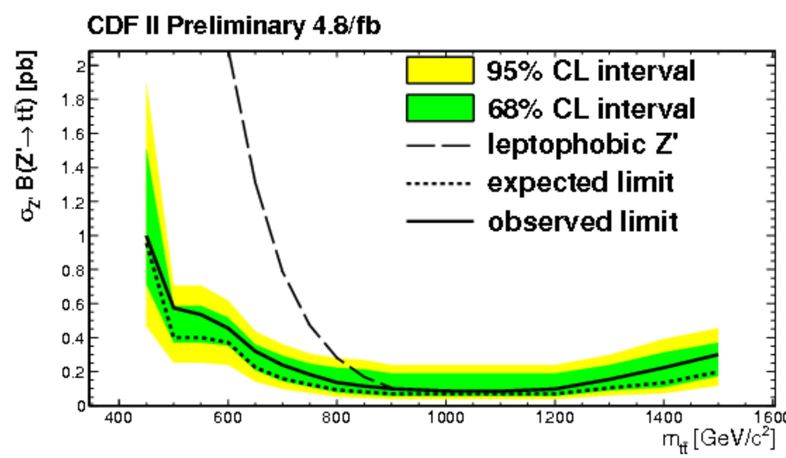
Boosted Top Quarks

- **Boosted top quarks a signature for several new physics models**
  - **Typically looking for resonances that decay to top-antitop pairs**
  - **Searches have focused on "resolved final states"**
    - Lepton+jets with b-tagging
    - Best limit is 70 fb at m_{tt}~1 TeV
    - Acceptance is 3.6%
  - **Limited by acceptance and production rate**
    - Exclude leptophilic Z' < 900 GeV/c^2
- **Our focus has been on unresolved final states**

CDF

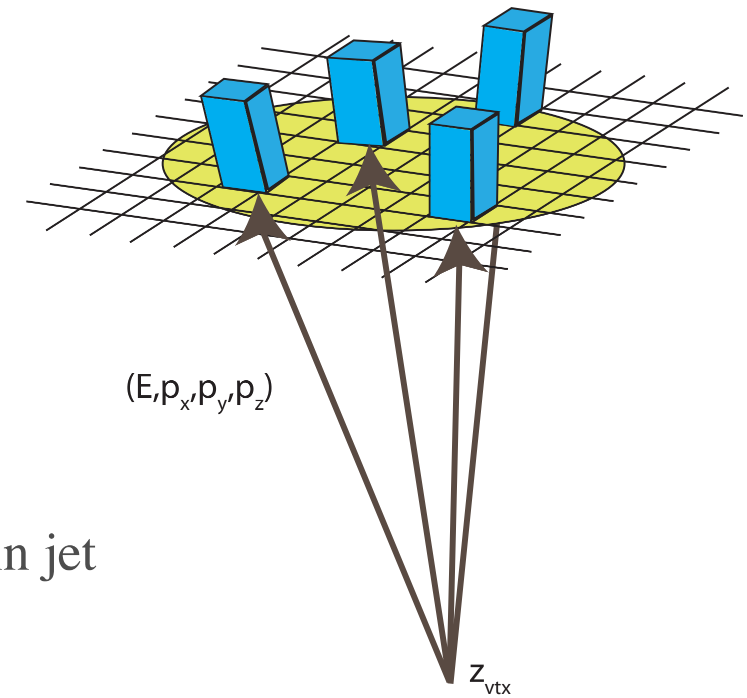


CDF, PRD 78, 052006 (2008)



Strategy for Analysis

- **Select high p_T jets in CDF central calorimeter**
  - **Use tower segmentation to measure jet mass**
    - Confirm with tracking information
  - **Employ standard "e-scheme" for mass calculation**
    - 4-vector sum over massless towers in jet
    - Four vector sum gives (E, p_x, p_y, p_z)
- **Employ Midpoint cone jets**
  - **Best understood in CDF II context**
  - **Compare results with anti-k_T and Midpoint with "search cones" (Midpoint/SC)**



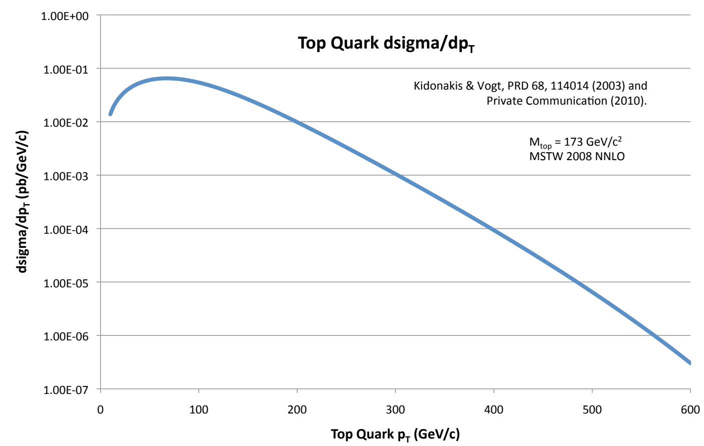
N.B. CDF central towers are $\Delta\eta \times \Delta\phi \sim 0.11 \times 0.26$

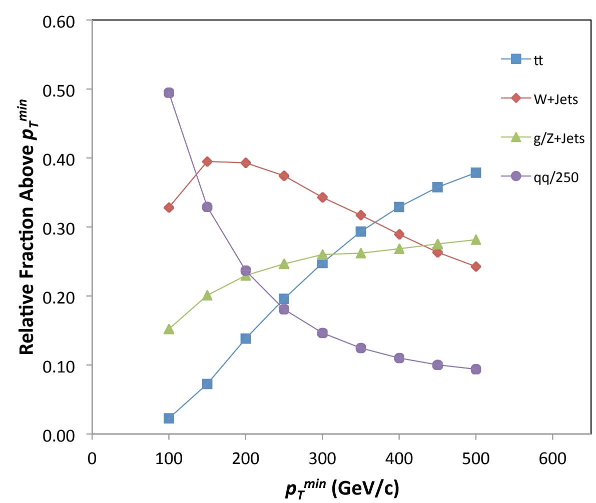
Boosted Objects at Tevatron

- **SM sources for high-p_T objects calculable**
  - **Dominated by light q & gluons**
  - **Need x250 rejection to observe other sources**

- **Other sources:**
  - **Fraction of top quarks ~1.5% for $p_T > 400$ GeV/c**
    - Total rate 4.45±0.5 fb (Kidonakis & Vogt)
    - PYTHIA 6.216 rate is 6.4 fb (scaling total cross section to measured world average)
  - **Expect W/Z production of similar order**

Kidonakis & Vogt, PRD 68, 114014 (2003)





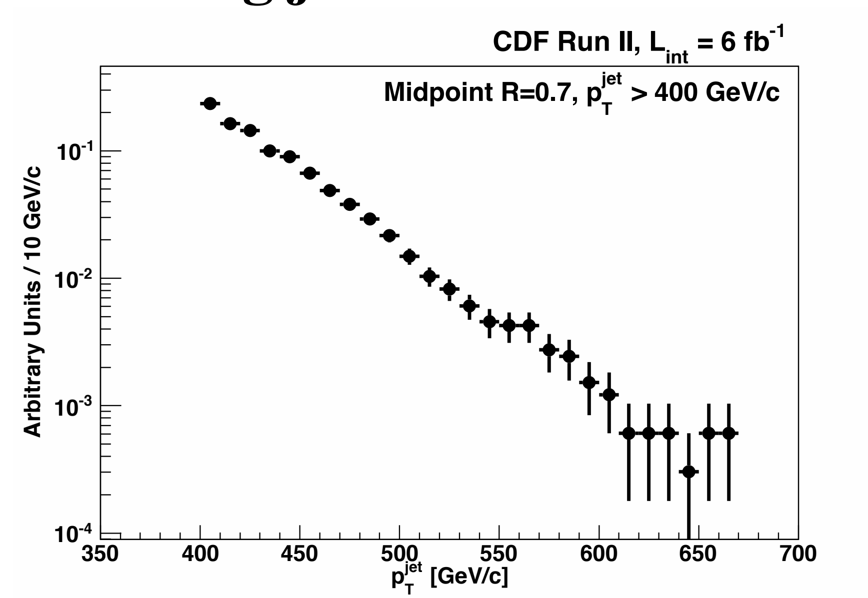
PYTHIA 6.4 Calculation

Data Selection

- **Analyzed inclusive jet sample**
  - **Trigger requires $E_T^{jet} > 100$ GeV**
  - **Analyzed 5.95 fb^{-1} sample**
- **Selected data with focus on high p_T objects**
  - **Kept any event with**
    - Jet with $p_T > 300$ GeV/c and $|\eta| < 0.7$
    - Used cones of R=0.4, 0.7 and 1.0
- **Processed 76M events**
  - **Selected subsample with**
    - $p_T > 400$ GeV/c
    - $|\eta| \in (0.1, 0.7)$
- **Performed cleaning cuts**

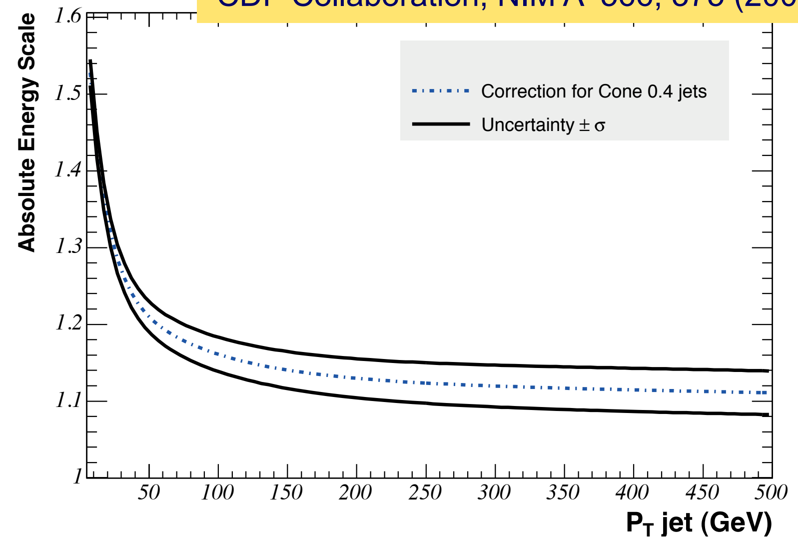
$$S_{MET} \equiv E_T^{MISS} \Big/ \sqrt{\sum_{i\ towers} E_T^i}$$

  - **Event vertex, jet quality and loose S_{MET} (< 14)**
- **Resulted in 2700 events using jets with R=0.7**



Jet Mass Corrections

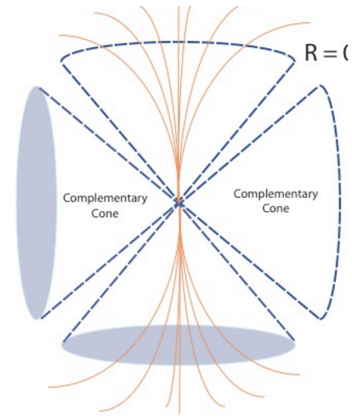
CDF Collaboration, NIM A 566, 375 (2006)

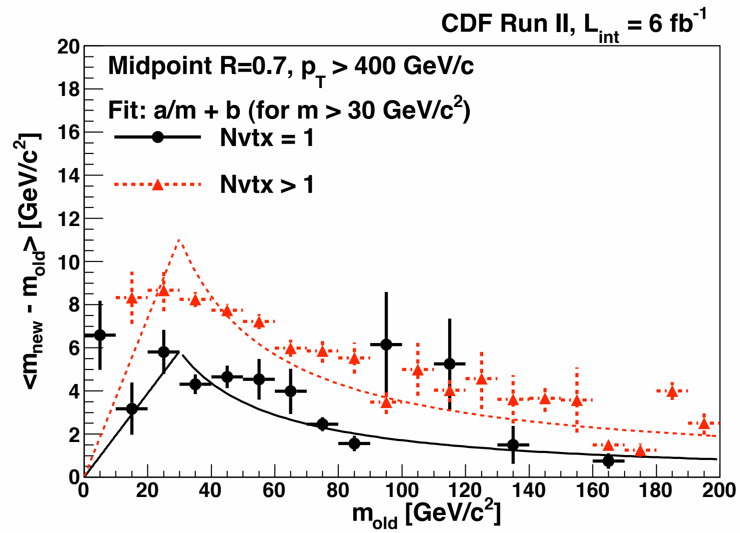


- **Corrected jet mass using standard jet corrections**
  - **Further correction needed for multiple interactions (MI)**
  - **Use N_{vtx}=1 and N_{vtx}>1 events to determine MI effect**
- **Investigated other effects:**
  - **Effect of calorimeter inhomogeneity at η=0**
    - Varied pseudorapidity window – no significant changes in mass
  - **Calorimeter segmentation and jet recombination**
    - Varied position of towers (especially azimuth) and corrections for geometry
  - **Calorimeter response across face of jet**
    - Detailed study of tracking/calorimeter response in data and MC/detector simulation
  - **Jet energy scale vs algorithm (Midpoint, Midpoint/SC, anti-k_T)**
    - Saw < 1 % difference

Effects of MI and UE

- **Additional contribution from**
  - **Underlying Event (UE)**
  - **Multiple Interactions (MI)**
    - Average # interactions ~3/crossing
- **Looked at purely dijet events**
  - **Defined cones (same size as jet) at 90º in azimuth (same η)**
  - **Took towers in cones, and added to leading jet in event**
    - Mass shift, on average, is same shift coming from UE and MI
- **Separately measure N_{vtx}=1 events**
  - **Gives UE correction separately**



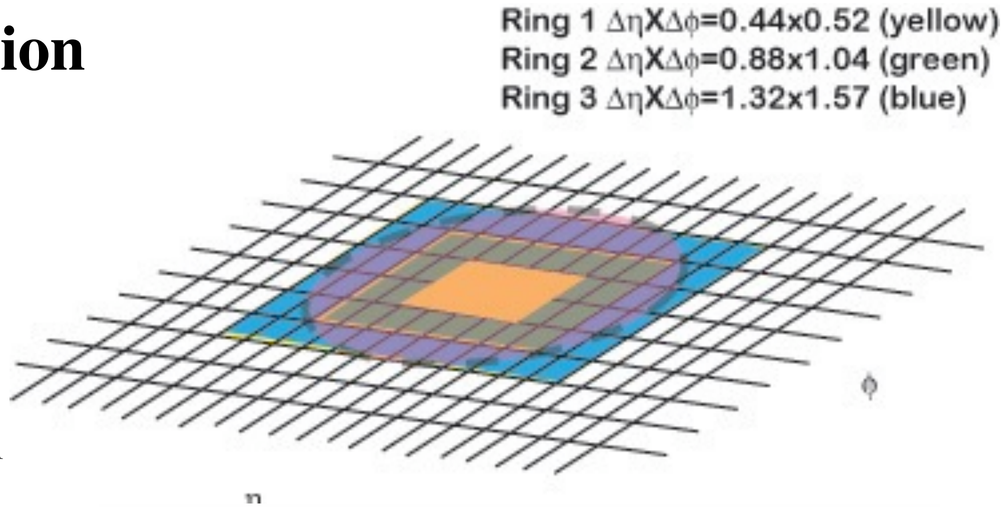


R. Alon et al., arXiv:1101.3002

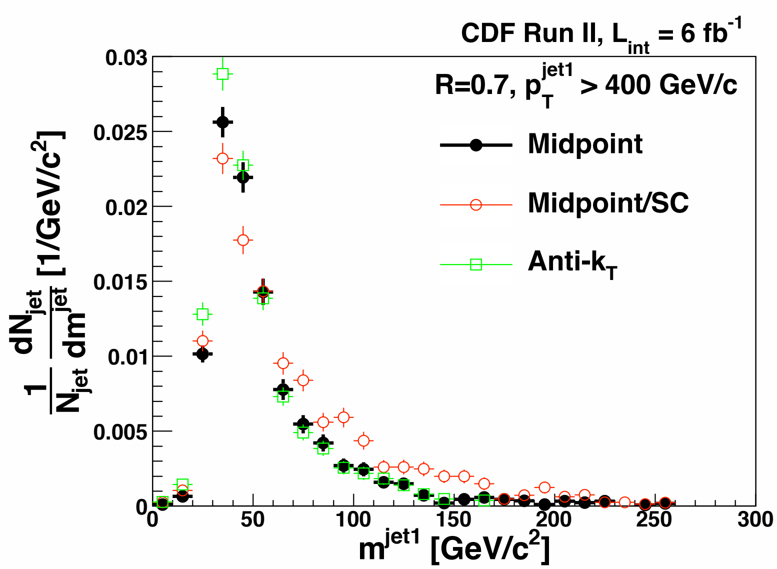
Correction scales as R^4

Inter-Jet Energy Calibration

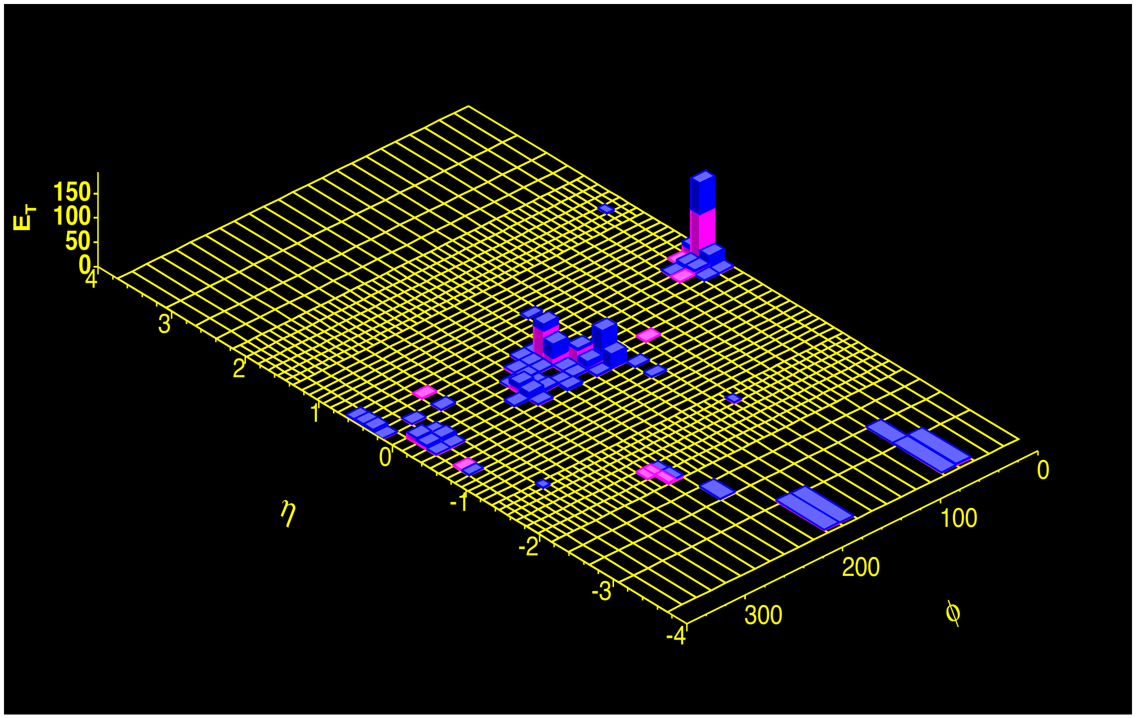
- **Jet mass arises from deposition of varying energy per tower**
  - **Performed study to compare momentum flow vs calorimeter energy internal to jet**
    - Defined 3 rings and compared observed p_T/E_T with simulation



- **Resulted in constraints on calorimeter relative response**
  - **At m^{jet}=60 GeV/c^2, Δm^{jet}=1 GeV/c^2**
  - **At m^{jet}=120 GeV/c^2, Δm^{jet}=10 GeV/c^2**
- **Largest source of systematic uncertainty**

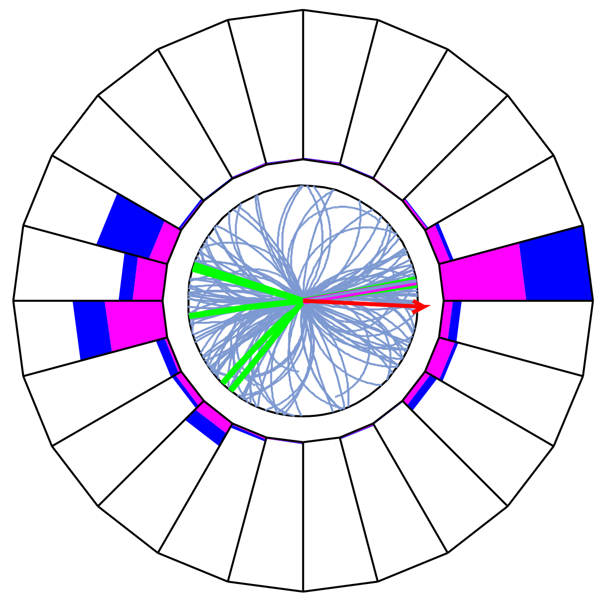


Typical Event



Run 286857 Event 79179

p_T	ϕ	m^{jet}	τ_{-2}	Pf
387	-3.11	175	0.024	0.66
344	0.09	113	0.019	0.40

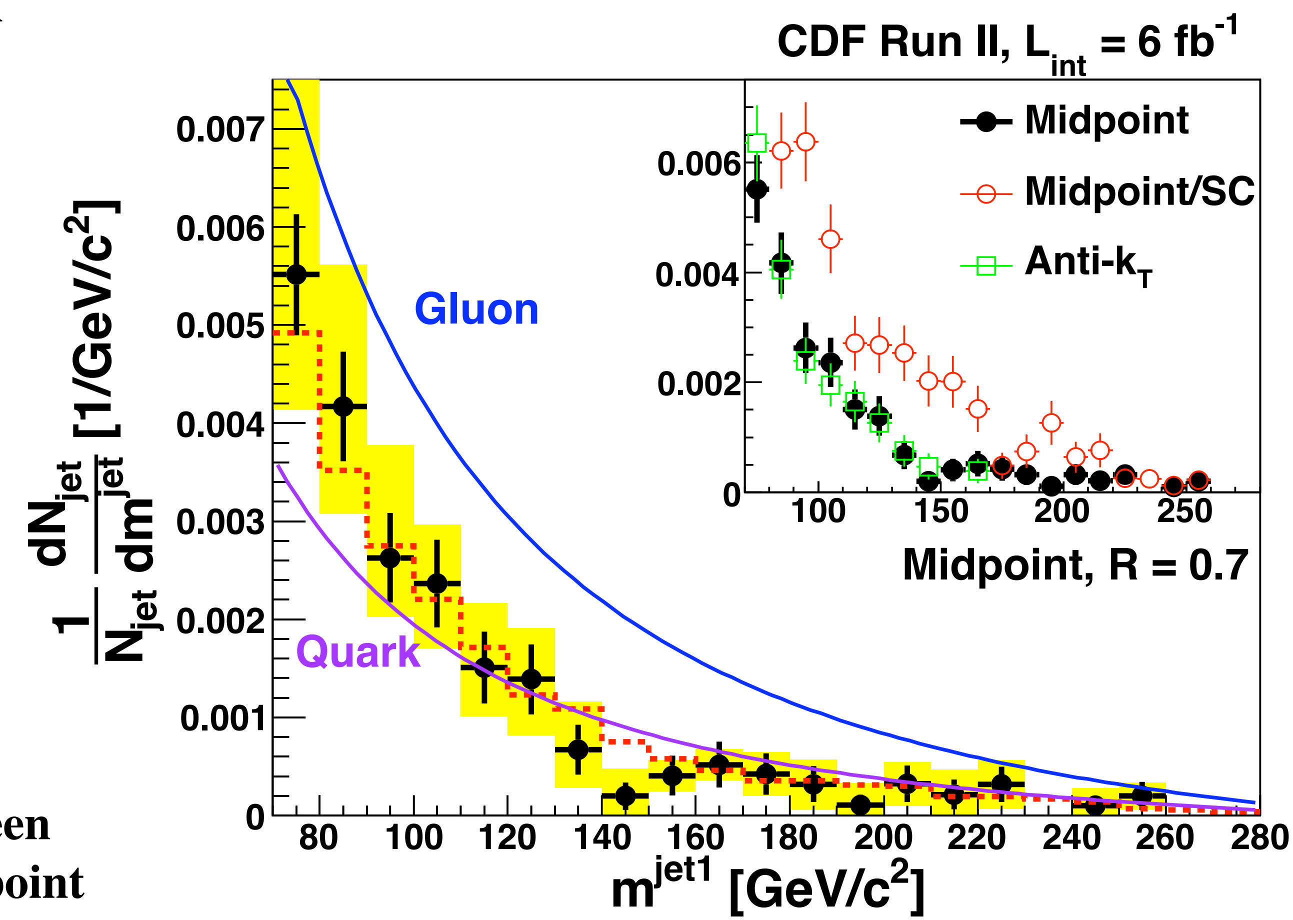


Typical QCD configuration:

- Dijet with back-to-back recoil
- Recoil jet less massive

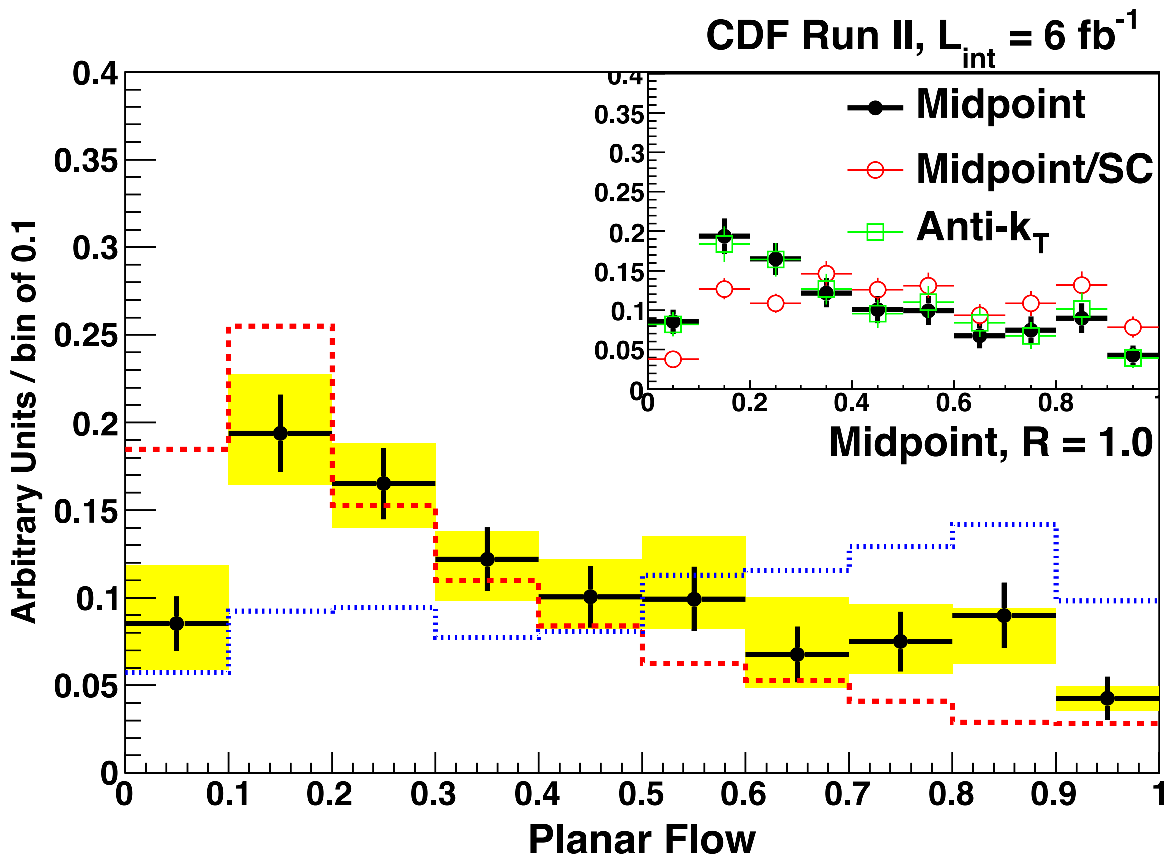
Jet Substructure – Mass

- **Massive jet**
  - Leading jets with $m_{jet} > 70$ GeV/c^2
  - Perform an "unfolding" correction
- **Agreement consistent with quark jets**
  - Expect ~85% of jets to be quark-initiated
  - No significant differences between anti-k_T and Midpoint algorithms



Jet Substructure – Planar Flow

- **Planar Flow is also IR-safe**
  - Low Pf -> two-body kinematics
  - Not strongly correlated to m^{jet} for high mass
- **Consistent with QCD predictions**
  - See the expected low Pf peak
  - Contrasts with top quark jets – larger planar flow

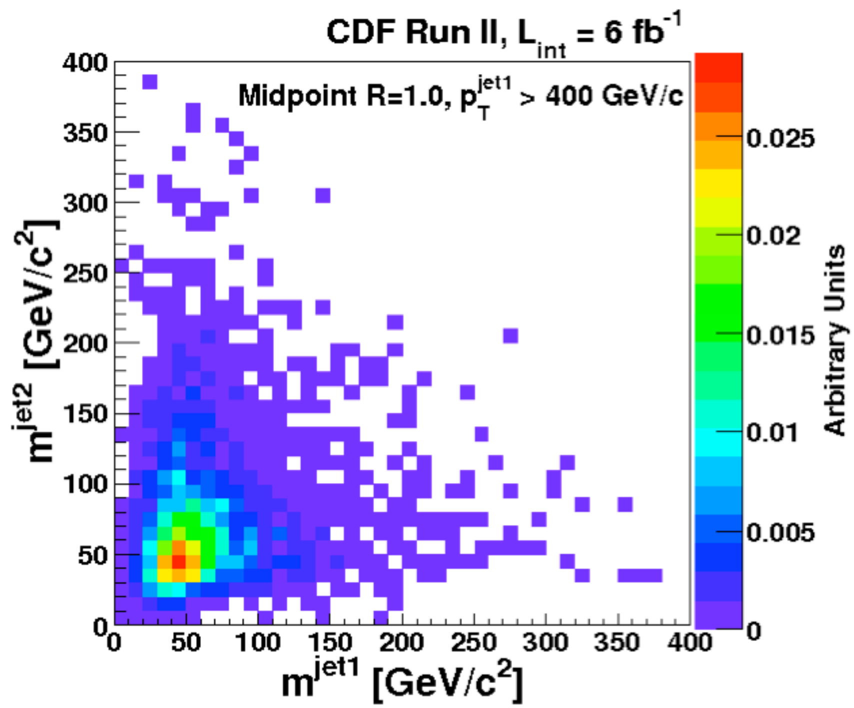


$130 < m_{jet} < 210$ GeV/c^2

Summary of Substructure Studies

- **Results show:**
  - **High p_T jets look like QCD light quark jets**
    - m^{jet} good discriminant
    - 1.4±0.3% of QCD jets have $m^{jet} > 140$ GeV/c^2
  - **Internal structure looks "two-body"**
    - Angularity & planar flow
  - **pQCD gives good description of mjet**
    - Other substructure measures well-modelled with PYTHIA

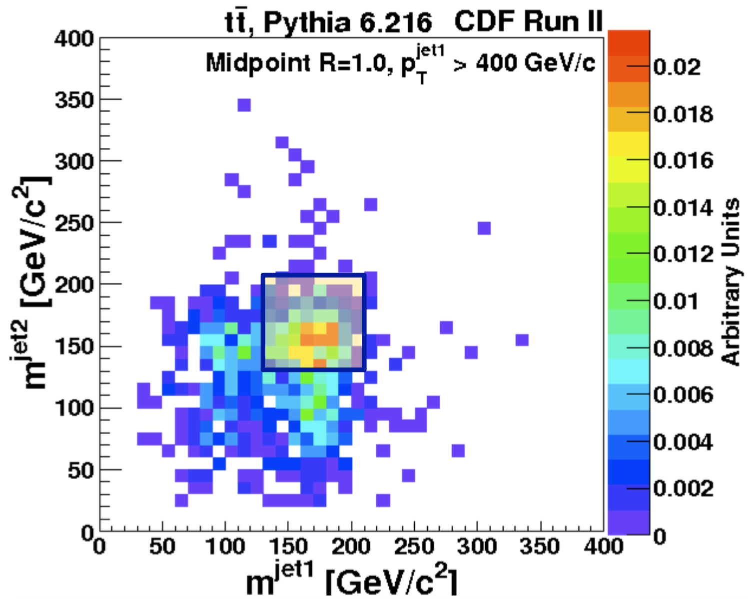
- **Jet masses are largely uncorrelated**
  - **Recoil jet doesn't know about leading m^{jet}**

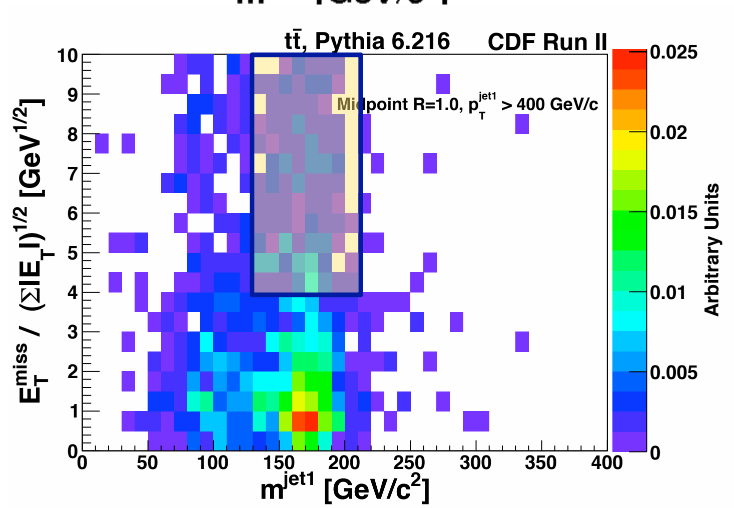


Strategies for Boosted Top

- **Two topologies:**
  1. **All hadronic ("1+1")**
     - Two massive jets recoiling (ε ~11%)
  2. **Semi-leptonic decay ("SL")**
     - Require $S_{MET} > 4$ (ε ~ 7%)
- **MC predicts ~0.8 fb**
  - **Divided 60:40 between topologies**
    - Highest efficiency channel for top (~18%)
  - **Important handles for background:**
    - masses of QCD di-jets not correlated
    - Jet mass and S_{MET} not correlated

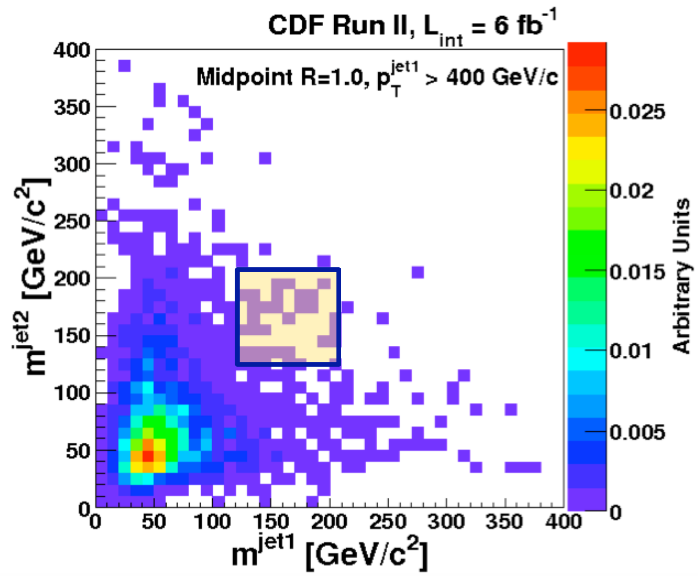
γ ~ 2.5

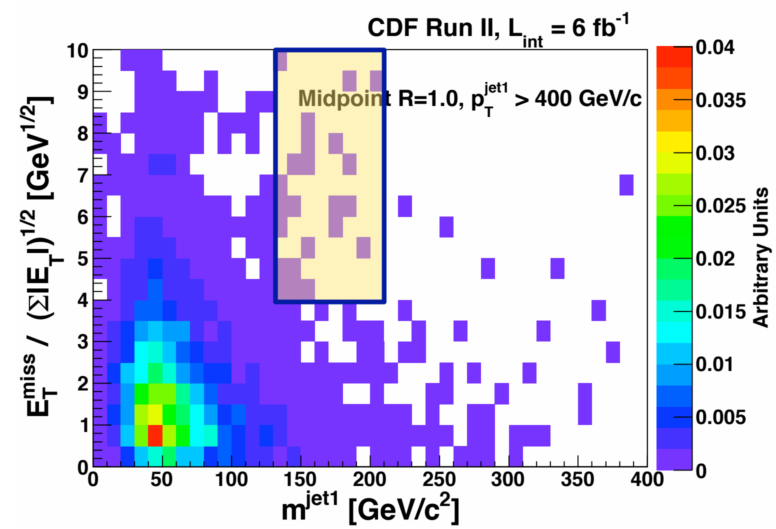




Selection Requirements

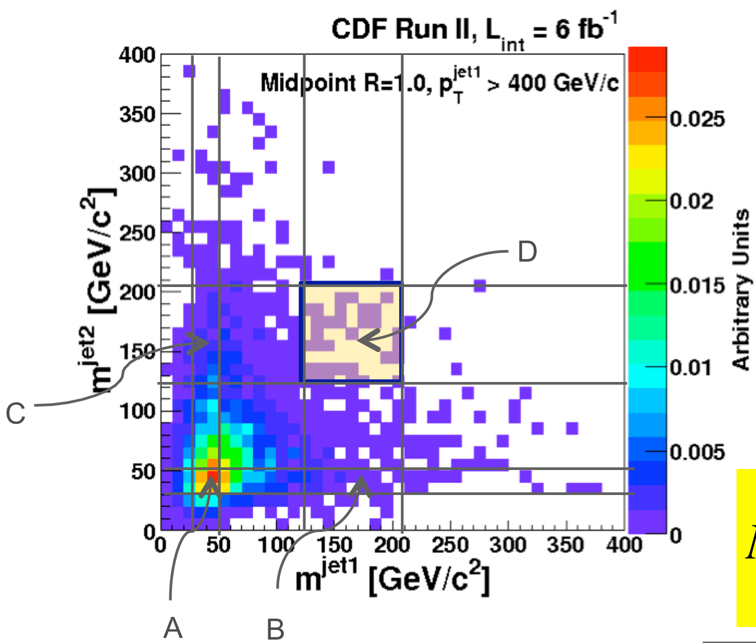
- **Keep selection simple**
  - Focus on two separate channels
- **All Hadronic Top (1+1)**
  - Require 2 jets with $130 < m^{jet} < 210$ GeV/c^2
  - Require $S_{MET} < 4$
  - Estimate background using "ABCD" technique
- **Semi-leptonic top (SL)**
  - Require $4 > S_{MET} > 10$
  - Require 1 jet with $130 < m^{jet} < 210$ GeV/c^2
  - Estimate background using "ABCD" technique





"Simple" Counting of 1+1

- With R=1.0 cones, m^{jet1} and m^{jet2} are equally powerful
  - Use jet mass (130,210) GeV/c^2 to define ttbar candidates
  - Expect 3.0±0.8 top quark events to populate this region



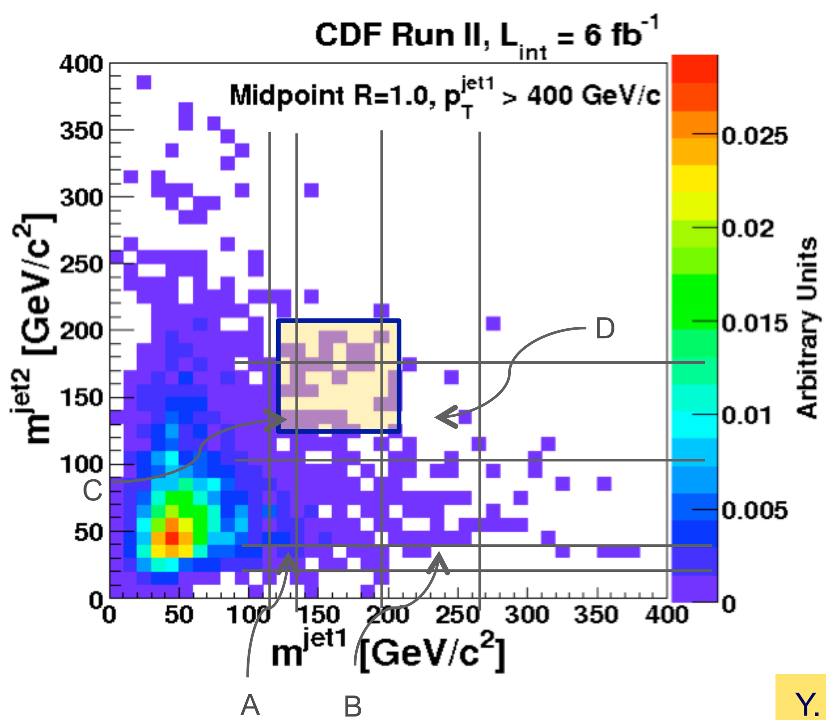
- Employ data to estimate backgrounds
  - Define mass windows
    $m^{jet} \in (130,210)$ GeV/c^2
    $m^{jet} \in (30,50)$ GeV/c^2
    - Use fact that m^{jet} distributions uncorrelated for background
    - Signal is region D
    - In "1+1" sample, predict 13±2.4 (stat) bkgd events

$$N_D^{Pred} = N_C \left[\frac{N_B}{N_A} \right]$$

- Observe N_D=32 events

Investigated m^{jet} Correlations

- We have been assuming that m^{jet1} and m^{jet2} are uncorrelated
  - Recent MC studies have shown this to be not exact



- NLO effects increase rate of two massive QCD jets
  - Quantified by defining R_{mass}

$$R_{mass} \equiv \left[\frac{N_C N_B}{N_A N_D}\right]$$

$$N_D^{pred} = \left[\frac{N_C N_B}{N_A R_{mass}}\right]$$

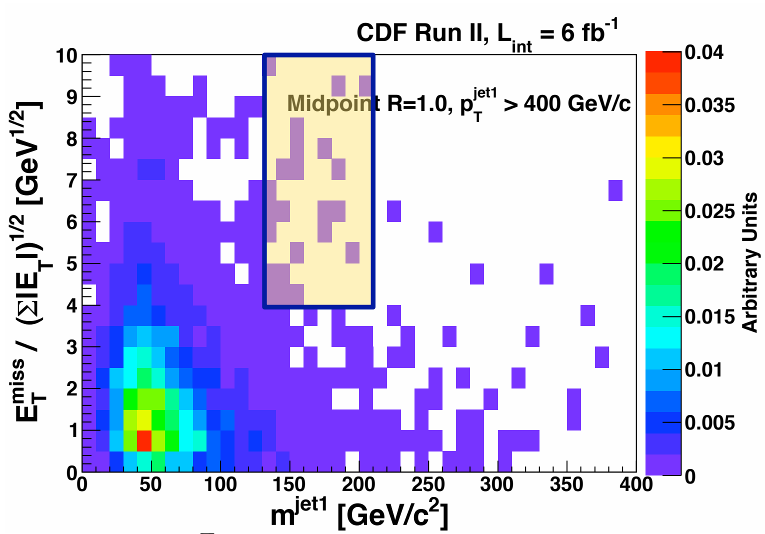
  - POWHEG: Rmass=0.89±0.03

MC tools	Matching	$R_{\rm mass}$
Sherpa	Yes	0.88 ± 0.03
MadGraph	Yes	0.86 ± 0.04
MadGraph	No	0.76 ± 0.04
Herwig	No	0.86 ± 0.02

Y. Eschel et al., arXiv:1101.2898

"Simple" Counting for SL

- **In case of recoil** semileptonic **top, use m^{jet1} and S_{MET}**
  - **Assumption is the S_{MET} and m^{jet1} are uncorrelated**
  - **Expect 1.9±0.5 top quark events to populate this region**



A B

- **Employ data to estimate backgrounds**
  - **Use regions $m^{jet1} \in (30,50)$ & (130,210) GeV/c^2**
  - **$S_{MET} \in (2,3)$ & $S_{MET} \in (4,10)$**
    - **In "SL" sample, predict 31±8 (stat) bkgd events**

- **Observe N_D=26 events**

Region	m^{jet1} (GeV/c^2)	S_{MET} ($\sqrt{GeV/c^2}$)	Data (Events)	MC (Events)
A	(30,50)	(2,3)	256	0.01
B	(130,210)	(2,3)	42	1.07
C	(30,50)	(4,10)	191	0.03
D (signal)	(130,210)	(4,10)	26	1.90
Predicted QCD in D			31.3 ± 8.1	

Uncertainties

- Background uncertainty (±10.2 GeV/c^2 jet mass scale)
  - ±30% uncertainty
- Uncertainties on top efficiency (SM production)
  - Primarily jet energy scale of ±3% on pT -> ±25% on σ
- Background statistics
  - ±11% from counting
- Luminosity
  - ±6% on integrated luminosity
- m^{top} uncertainty (±2 GeV/c^2)
  - ±0.3%
- Overall uncertainties added in quadrature
  - ±41% overall
- Incorporated into upper limit calculation
- Use a CL_s frequentist method
  - Marginalize nuisance parameters
  - Same as used in Higgs and single top searches

Top Quark Cross Section Limit

- **Assume we observe signal + background**
  - **Set upper limit on SM production σ for top quark with $p_T > 400$ GeV/c**
- **Observe 58 events with 44+/-8 background**
  - **Calculate 95% CL upper limit using CL_s method**
    - Systematic uncertainties incorporated a la CDF 8128 (T. Junk)
    - $N_{LIM} = 43.3$ events
  - **Efficiency from MC**
    - 1+1: 11.1%
    - SL: 7.0%

- **Upper limit on cross section for $p_T > 400$ GeV/c**

$$\sigma_{95\%} = \frac{N_{LIM}}{\int L\,dt\,\varepsilon}$$

$$= \frac{43.3}{(5.95)(0.182)} = 40\text{ fb}$$

- **Can also set limit on 1+1 only**
  - **Assume massive ($m \sim m_{top}$) object, pair-produced, decaying hadronically**
  - **Include SM top as background**

$$\sigma_{95\%} = \frac{N_{LIM}}{\int L\,dt\,\varepsilon}$$

$$= \frac{30.2}{(5.95)(0.254)} = 20\text{ fb}$$

Also ~3σ excess above SM top

Conclusions

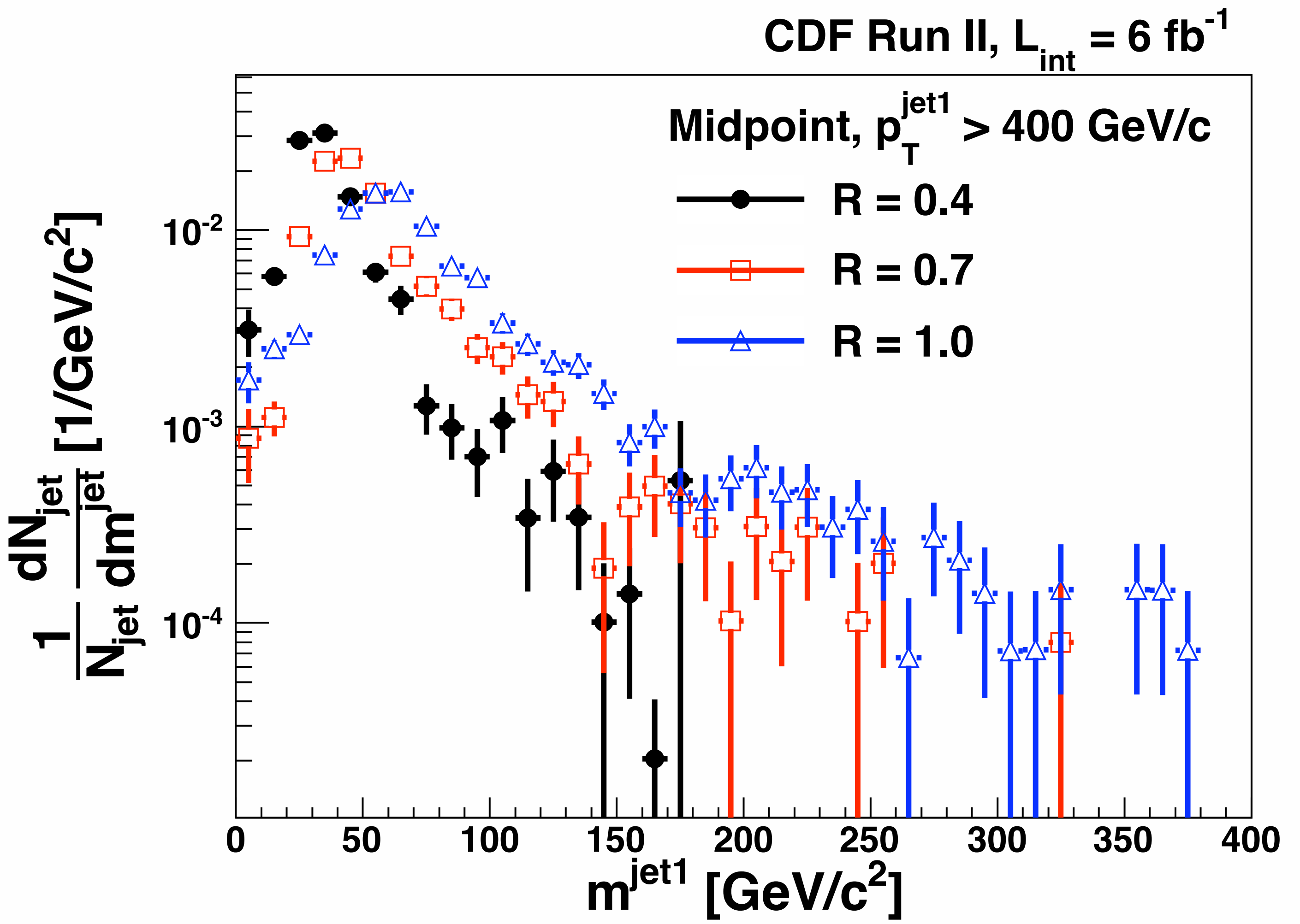
- **Search for boosted top at Tevatron close to SM rate**
  - **Achieve**

    $$S/\sqrt{B} \approx 0.75$$
  - **Set $\sigma < 40$ fb at 95% CL**
  - **Limited by statistics**
- **Doesn't take advantage of substructure (aside from m^{jet})**
  - **E.g., planar flow cut > 0.5 improves S/N by ~1.5**
  - **And haven't used**
    - B-tagging
    - For SL, look for isolated charge track
- **Next steps**
  - **At Tevatron, can improve statistics by x2**
  - **Tantalizing close to SM**
  - **Ultimately limited by rate**
- **Real focus are LHC expts**
  - **Now recorded sample with similar # of boosted SM ttbar**
    - But QCD backgrounds are larger
  - **Jet substructure is clearly essential tool**
    - Fully characterize QCD jets
    - Understand what the best tools are

BACKUP SLIDES

Comparison with Cone Size

- Compare
  - R=0.4
  - R=0.7
  - R=1.0



Jet Algorithms

- **Cone algorithms used for most Tevatron studies**
  - **Long history – quite separate from e^+e^- work**
  - **JetClu was CDF reference**
    - Required "seed" to initiate
    - Significant IRC sensitivity
- **Midpoint developed to reduce IRC sensitivity**
  - **Use seeds, but then recluster with seeds "midway" between all jets**

Use Fastjet Framework!

M. Cacciari, G.P. Salam and G. Soyez, Phys. Lett. B641, 57 (2006) [hep-ph/0512210].

- **Cone algorithms had "dark tower" problem**
  - **Unclustered energy due to split/merge/iteration procedure**
  - **Proposed solution: Midpoint with "search cones"**
    - Find jets with cone size R/2
    - Fix jet direction, cluster with size R
  - **Midpoint/SC was used for various studies 2006-2008**
- **Anti-k_T algorithm developed**
  - **No IR sensitivity**
  - **Still retained many of the benefits of a "cone" algorithm**

MI/UE Corrections

- Looked at how to make MI correction in a variety of ways
  - Looked at mass corrections event-by-event
  - But statistical fluctuations large, event-to-event
  - Chose to develop a parametrized correction
- Note that:

$$\delta m^{jet} \simeq \frac{E_{tower} E_{jet} \Delta R}{m^{jet}}$$

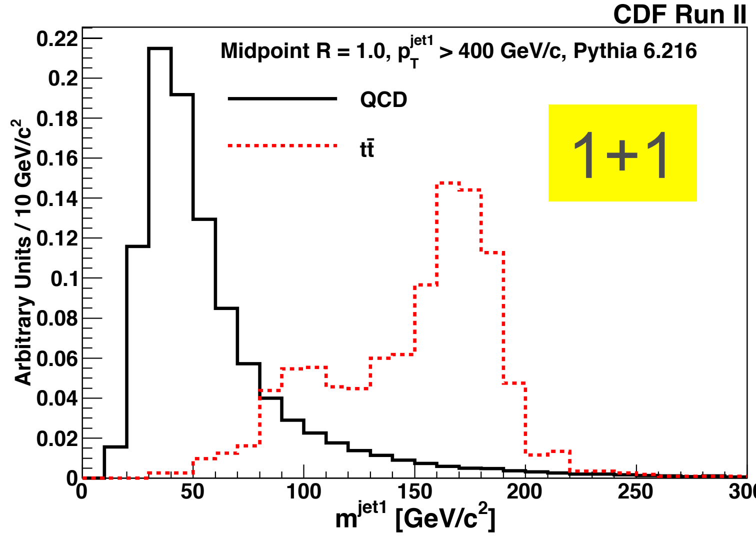
- Expect MI correction to scale with R^4:
  - Exactly what we see when comparing R=0.4 and R=0.7
- PYTHIA UE agrees well with data – same UE mass correction
- Use that to scale corrections for R=1.0
  - Method doesn't work with larger cone because of overlap

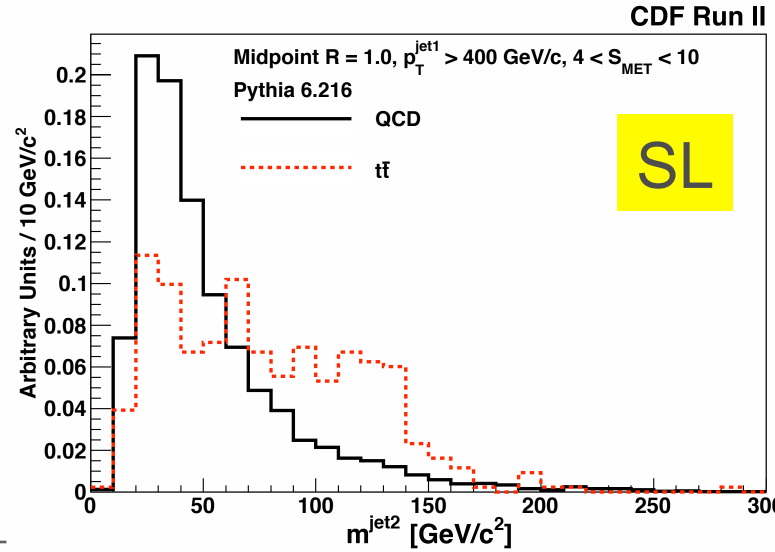
Internal Jet Energy Scale

- Overall jet energy scale known to 3%
  - The relative energy scale between rings known to 10-20%, depending on ring
  - Use this to constrain how far energy scale can shift
- Do first for $m^{jet} \sim 60$ GeV/c^2 – use average jet profile
  - Extract from that a limit on how much "Ring 1" energy scale can be off - ± 6%
  - Then do the same for mjet ~ 120 GeV/c^2
- Resulting systematic uncertainty is 9.6 GeV/c^2
  - Conservative estimate – used a very broad energy profile
    - No localized substructure assumed
- Take this as systematic uncertainty
  - Could constrain it better using single particle response
  - Note that fixed cone size is an advantage here

Reconstruction of Top

- **Leading jet in ttbar events has clear top mass peak**
  - **All events between 70 and 210 GeV/c^2 for R=1.0**
  - **See evidence of W peak**
    - B quark jet presumably nearby in those cases
  - **Clear that higher mass cut gives greater QCD rejection**
    - But also start to lose efficiency
  - **S_{MET} cut effectively identifies semi-leptonic decays (8%)**
- **B tagging not used**
  - **Can estimate mis-tags using data -> ~0.05%/jet**
  - **But large uncertainty in tagging efficiency in high pT jets**





CDF Collaboration

Background Calculations

- Background calculations used "ABCD" technique

- SL

Region	m^{jet1} ($\mathrm{GeV/c^2}$)	S_{MET} ($\sqrt{GeV/c^2}$)	Data (Events)	MC (Events)
A	$(30,50)$	$(2,3)$	256	0.01
B	$(130,210)$	$(2,3)$	42	1.07
C	$(30,50)$	$(4,10)$	191	0.03
D (signal)	$(130,210)$	$(4,10)$	26	1.90
Predicted QCD in D			31.3 ± 8.1	

- 1+1

Region	m^{jet1} ($\mathrm{GeV/c^2}$)	m^{jet2} ($\mathrm{GeV/c^2}$)	Data (Events)	$t\bar{t}$ MC (Events)
A	$(30,50)$	$(30,50)$	370	0.00
B	$(130,210)$	$(30,50)$	47	0.08
C	$(30,50)$	$(130,210)$	102	0.01
D (signal)	$(130,210)$	$(130,210)$	32	3.03
Predicted QCD in D			13.0 ± 2.4	